

Detection of anatomical changes in lung cancer patients with 2D time-integrated, 2D time-resolved and 3D time-integrated portal dosimetry

Citation for published version (APA):

Wolfs, C. J. A., Bras, M. G., Schyns, L. E. J. R., Nijsten, S. M. J. J. G., van Elmpt, W., Scheib, S. G., Baltes, C., Podesta, M., & Verhaegen, F. (2017). Detection of anatomical changes in lung cancer patients with 2D time-integrated, 2D time-resolved and 3D time-integrated portal dosimetry: a simulation study. *Physics in Medicine and Biology*, 62(15). <https://doi.org/10.1088/1361-6560/aa7730>

Document status and date:

Published: 07/08/2017

DOI:

[10.1088/1361-6560/aa7730](https://doi.org/10.1088/1361-6560/aa7730)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 19 Apr. 2024

PAPER

Detection of anatomical changes in lung cancer patients with 2D time-integrated, 2D time-resolved and 3D time-integrated portal dosimetry: a simulation study

To cite this article: Cecile J A Wolfs *et al* 2017 *Phys. Med. Biol.* **62** 6044

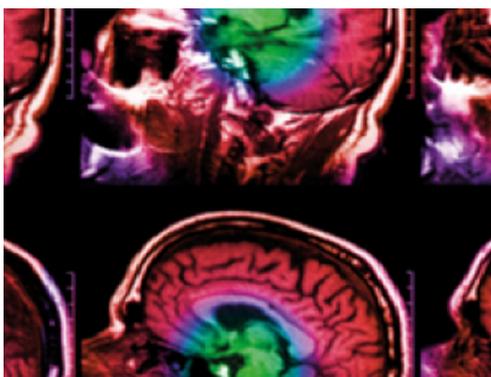
View the [article online](#) for updates and enhancements.

You may also like

- [Domain adaptation-based deep feature learning method with a mixture of distance measures for bearing fault diagnosis](#)
Kaibo Zhou, Guannan Cao, Kaifeng Zhang *et al.*
- [Two-dimensional analysis of the nuclear relaxation function in the time domain: the program CracSpin](#)
W P Węglarz and H Haraczuk
- [Fetal ECG extraction using short time Fourier transform and generative adversarial networks](#)
Wei Zhong and Weibin Zhao

Recent citations

- [Identification of treatment error types for lung cancer patients using convolutional neural networks and EPID dosimetry](#)
Cecile J.A. Wolfs *et al*
- [Katia Parodi and Christian Thieke](#)
- [Three-dimensional dose evaluation in breast cancer patients to define decision criteria for adaptive radiotherapy](#)
Catharina M. L. Zegers *et al*



IPEM | IOP

Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics,
biomedical engineering and related subjects.

Start exploring the collection—download the
first chapter of every title for free.

Detection of anatomical changes in lung cancer patients with 2D time-integrated, 2D time-resolved and 3D time-integrated portal dosimetry: a simulation study

Cecile J A Wolfs^{1,3}, Mariana G Brás^{1,3}, Lotte E J R Schyns¹, Sebastiaan M J J G Nijsten¹, Wouter van Elmpt¹, Stefan G Scheib², Christof Baltes², Mark Podesta¹ and Frank Verhaegen¹

¹ Department of Radiation Oncology (MAASTRO), GROW—School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, the Netherlands

² Varian Medical Systems iLab GmbH, Research & Development, Dättwil Baden, Switzerland

E-mail: frank.verhaegen@maastro.nl

Received 10 February 2017, revised 8 May 2017

Accepted for publication 5 June 2017

Published 12 July 2017



Abstract

The aim of this work is to assess the performance of 2D time-integrated (2D-TI), 2D time-resolved (2D-TR) and 3D time-integrated (3D-TI) portal dosimetry in detecting dose discrepancies between the planned and (simulated) delivered dose caused by simulated changes in the anatomy of lung cancer patients.

For six lung cancer patients, tumor shift, tumor regression and pleural effusion are simulated by modifying their CT images. Based on the modified CT images, time-integrated (TI) and time-resolved (TR) portal dose images (PDIs) are simulated and 3D-TI doses are calculated. The modified and original PDIs and 3D doses are compared by a gamma analysis with various gamma criteria. Furthermore, the difference in the $D_{95\%}$ ($\Delta D_{95\%}$) of the GTV is calculated and used as a gold standard. The correlation between the gamma fail rate and the $\Delta D_{95\%}$ is investigated, as well the sensitivity and specificity of all combinations of portal dosimetry method, gamma criteria and gamma fail rate threshold.

On the individual patient level, there is a correlation between the gamma fail rate and the $\Delta D_{95\%}$, which cannot be found at the group level. The

³These authors contributed equally to this work.

sensitivity and specificity analysis showed that there is not one combination of portal dosimetry method, gamma criteria and gamma fail rate threshold that can detect all simulated anatomical changes.

This work shows that it will be more beneficial to relate portal dosimetry and DVH analysis on the patient level, rather than trying to quantify a relationship for a group of patients. With regards to optimizing sensitivity and specificity, different combinations of portal dosimetry method, gamma criteria and gamma fail rate should be used to optimally detect certain types of anatomical changes.

Keywords: portal dosimetry, adaptive radiotherapy, anatomical changes, gamma analysis, DVH metrics

(Some figures may appear in colour only in the online journal)

1. Introduction

The development of complex beam delivery techniques in external beam radiotherapy, such as intensity modulated radiotherapy (IMRT) and volumetric modulated arc therapy (VMAT) (Bortfeld 2006, Teoh *et al* 2014), has increased the need for accurate verification of the dose delivery. Due to the highly dynamic nature of VMAT, quality assurance is more difficult to perform than for conventional static radiation delivery techniques. The electronic portal imaging device (EPID) has been successfully used as a tool for performing dose verification for IMRT and VMAT; initially in a time-integrated manner, but more recently also in a time-resolved fashion (van Elmpt *et al* 2008, Liu *et al* 2013, Woodruff *et al* 2013, Podesta *et al* 2014a, Spreeuw *et al* 2016).

Besides being able to detect errors related to beam delivery or plan transfer in a pre-treatment setting, transit EPID dosimetry has been shown to be valuable in detecting errors at the patient level during treatment (Nijsten *et al* 2007, Mans *et al* 2010, Persoon *et al* 2012, 2013, 2015b). Anatomical changes occur commonly during the course of treatment, and can cause discrepancies between the planned and actually delivered dose distributions that may need corrective actions (Yan *et al* 1997, McDermott *et al* 2006). It is therefore important to perform dose verification during the patient's course of treatment to detect these errors and make the necessary adjustments to ensure treatment quality. Cone-beam CT imaging can be used to detect anatomical changes as well. However, this modality does not allow to evaluate directly discrepancies in the delivered dose due to anatomical changes. Furthermore, with cone-beam imaging additional dose is delivered to the patient, and additional scan time is needed. Transit portal dosimetry does not pose these disadvantages.

In this work, the performance of 2D time-integrated (2D-TI), 2D time-resolved (2D-TR) and 3D (*in vivo*) time-integrated (3D-TI) transit portal dosimetry are evaluated. In 2D, the dose distribution behind the patient at the level of the portal imager is evaluated, whereas in 3D the dose inside the patient is calculated (van Elmpt *et al* 2008). Furthermore, with time-integrated (TI) portal dosimetry, the dose is accumulated over a VMAT arc, while time-resolved (TR) portal dosimetry evaluates the dose distributions per VMAT segment (time between consecutive control points) (Bedford *et al* 2014, Podesta *et al* 2014a, Persoon *et al* 2015b).

The focus of this work is on detecting dose discrepancies between the planned and (simulated) delivered dose caused by simulated changes in the patient's anatomy. Simulating anatomical changes and the resulting dose distributions allows for a systematic, high resolution

analysis of the detection limits of portal dosimetry. The aim of this work is to determine these detection limits, and to discover trends that can ultimately aid in developing clinical decision models for dose guided adaptive radiotherapy.

2. Methods and materials

Tumor shift, tumor regression and pleural effusion are commonly occurring anatomical changes during treatment of lung cancer patients, and can cause errors in the delivered dose distributions (Britton *et al* 2007, Kwint *et al* 2014, Møller *et al* 2014). To study the ability of portal dosimetry to detect these errors, these three types of anatomical changes were simulated using planning CT images of six lung cancer patients treated at MAASTRO Clinic. All patients were treated with VMAT treatments consisting of two (half-)arcs, using 6 or 10 MV photon beams delivered by a Varian TrueBeam linear accelerator (Varian Medical Systems, Palo Alto (CA), United States). Some characteristics of the patients and their treatments are outlined in table 1.

2.1. Anatomical change simulation

Figure 1 illustrates the framework used in this work for simulating anatomical changes, simulating dose distributions and comparing the dose distributions. The simulation of anatomical changes was performed by modifying the patients' planning CT images with an in-house developed tool implemented in MATLAB (The Mathworks, Inc., Natick (MA), United States). All simulations and evaluations were based on the gross tumor volume (GTV), as this allowed for modifying and assessing a structure with visible boundaries rather than the invisible boundaries derived from extending the GTV to a clinical target volume (CTV) and planning target volume (PTV).

A baseline tumor shift was introduced by changing the position of the GTV. The voxels of the CT image where the tumor was located before shifting, were forced to a Hounsfield unit (HU) value of -750 , which was approximately the average of the HUs in the lungs of all patients. The shifts were simulated along the craniocaudal direction, as it has been shown previously by our group that shifts along this direction induce larger discrepancies (i.e. larger gamma fail rates) in measured portal dose distributions (Schyns *et al* 2016). Thirteen consecutive shifts of 0.3 cm were applied resulting in a range of shifts from 0.3 to 3.9 cm. These levels of severity were chosen because the voxel size of the CT in the craniocaudal direction was 0.3 cm, whereas the maximum shift ensured the effect of moving the GTV (almost) completely out of the PTV was taken into account.

Tumor regression was simulated by applying a centroid-based deformation to the GTV. As with the simulation of tumor shift, the HUs of the voxels where the tumor was before the regression were forced to -750 . The levels of regression simulated ranged from 10% to 90% of the original GTV volume, in steps of 10%.

In the case of pleural effusion, gravity will often cause the fluid to accumulate on the dorsal side of the pleural cavity, as the patient is in the supine position during treatment. Therefore, pleural effusion was simulated by creating a pleural effusion volume on the dorsal side of the lung the tumor was located in, and forcing the HUs in this volume to 0 (the value of water). The range of simulated pleural effusion levels was either from 0.5 to 4.0 cm in steps of 0.5 cm, or from 0.1 to 0.8 cm in steps of 0.1 cm, depending on whether the tumor was located at the ventral side of the lung or more towards the dorsal side, respectively.

Table 1. Characteristics of the six lung cancer patients and their treatments included in this study.

Patient and treatment characteristics	
Gender	
Male	3/6
Female	3/6
Age (mean (range))	75 (68–85)
Tumor location	
Left lung	3/6
Right lung	3/6
Upper lobe	3/6
Lower lobe	3/6
Middle lobe (right lung only)	0/3
Tumor volume (mean (range))	31.5 cm ³ (4.6–74.7) cm ³
Margins	
GTV to CTV	
5 mm	3/6
Anisotropic (~5 mm)	3/6
CTV to PTV (Individual margin based on internal target volume (ITV))	
5 mm	1/6
8 mm	2/6
10 mm	2/6
Anisotropic (~5 mm)	1/6
Dose prescription (dose per fraction (Gy)/number of fractions)	
1.5/30	1/6
1.8/41 or 43	2/6
2.75/24	2/6
7.5/8	1/6

After introducing anatomical changes of different magnitudes in each patient's planning CT, 2D portal dose images (PDIs), at a source-to-detector distance (SDD) of 150 cm, were simulated for the original planning CTs and for each modified CT, with an in-house developed portal dose prediction model (van Elmpt *et al* 2005). All PDI predictions were performed based on each patient's original RT plan. Furthermore, 3D doses in the patients were calculated with a Monte Carlo based dose reconstruction method (van Elmpt *et al* 2006). In 2D, both TI and TR PDIs were predicted, while in 3D only TI dose distributions were simulated.

The 2D PDIs and 3D dose distributions of the original planning CT and those of the modified CTs were compared using a global gamma analysis (Low *et al* 1998, Low and Dempsey 2003). For 2D-TR portal dosimetry a time-resolved gamma analysis was used (Podesta *et al* 2014b). As the data is simulated, there is perfect synchronization in time between the original and modified data, meaning that the time of interest parameter in this analysis was set to 0.

The most commonly used gamma criteria for dose difference and distance to agreement are (3%, 3 mm) (Low *et al* 2013). Yet, studies have indicated that these criteria are not sensitive enough to detect relevant dose discrepancies (Yan *et al* 2009, Stasi *et al* 2012, Heilemann *et al* 2013). To study the performance of various gamma criteria in detecting errors, six sets of gamma criteria were employed, namely (1%, 1 mm), (1.5%, 1.5 mm), (2%, 2 mm), (2.5%,

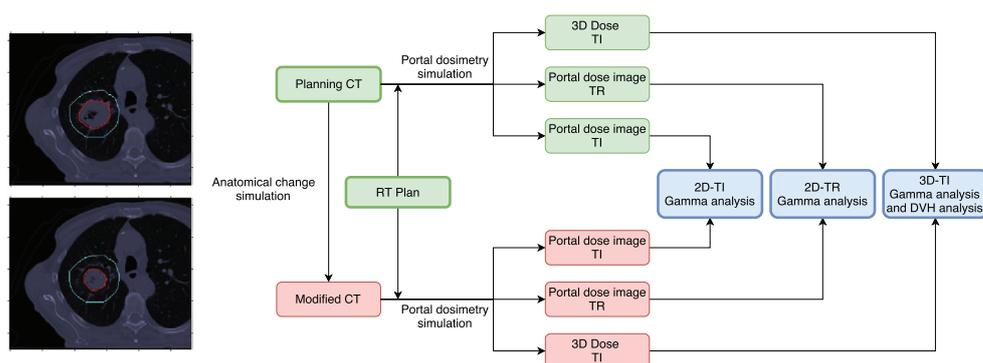


Figure 1. A flowchart of the process of simulating anatomical changes, simulating dose distributions and comparing the dose distributions used in this work. The bold outlined boxes on the left and right represent the input and output, respectively. The boxes in green depict all data not containing simulated anatomical changes, while those in red do contain (the results of) simulated anatomical changes. The blue boxes represent the comparisons between the simulated dose distributions. On the far left an example of simulated tumor regression (50%) is displayed, with the planning CT (top) and the modified CT (bottom). The GTV is delineated in red, the PTV (for reference, not modified) in blue.

2.5 mm), (3%, 3 mm) and (5%, 3 mm). The results of the global gamma analyses are expressed in terms of the gamma fail rate, which is the percentage of failing pixels (i.e. pixels with $|\gamma| > 1$) inside the radiation field in 2D (accumulated over the VMAT arc for TI and per segment for TR), and the percentage of failing voxels inside the GTV structure in 3D. To evaluate the performance of the three portal dosimetry methods (2D-TI, 2D-TR and 3D-TI) in detecting the simulated anatomical changes, it was first determined for which cases the gamma fail rate exceeded 10%, which is considered a clinically acceptable threshold for detecting errors in the delivered dose at our clinic. The analysis can be repeated with another value for this parameter.

2.2. Correlation between DVH metrics and gamma fail rate

In clinical practice, dose-volume histogram (DVH) analysis is an important part of treatment plan optimization and comparison (Fraass *et al* 1998). In this work, it was considered the gold standard to detect relevant dose discrepancies. For the design of decision support protocols for adaptive dose guided radiotherapy and to prevent subjective decision making, a quantitative relationship between gamma analysis results and DVH metrics should be established. Therefore, DVHs were generated for each 3D-TI dose distribution. The DVH metric considered in this work is the dose to 95% ($D_{95\%}$) of the GTV. The percentage of change in this metric in comparison to the original GTV was determined as follows:

$$\Delta D_{95\%} = \left| \frac{D_{95\%}(\text{modified GTV}) - D_{95\%}(\text{original GTV})}{D_{95\%}(\text{original GTV})} \right| \cdot 100$$

The correlation between the $\Delta D_{95\%}$ and the gamma fail rate was investigated for all three portal dosimetry methods and for all six sets of gamma criteria. The correlation analysis consisted of calculating Pearson's r for each patient separately, and for the six patients as a group. The correlation coefficient is interpreted as representing no linear relationship for r between 0 and ± 0.2 , a weak relationship for r between ± 0.2 and ± 0.5 , a moderate relationship for r between

± 0.5 and ± 0.8 and a strong relationship for r between ± 0.8 and 1 (Zou *et al* 2003). In the case of tumor shift, a saturation-growth curve $y = a \frac{x}{b+x}$ was fitted to the data, and this non-linear fit was assessed by the coefficient of determination R^2 .

2.3. Sensitivity and specificity analysis

Sensitivity and specificity analysis using receiver operating characteristic (ROC) curves is widely used to determine the performance of a dichotomous diagnostic test against a gold standard (Kumar and Indrayan 2011). In this case, the diagnostic tests were the portal dosimetry methods with their resulting gamma fail rates, which can be either positive (i.e. exceeding a certain threshold) or negative (i.e. not exceeding a certain threshold). As discussed in the previous section, DVH analysis, and in this case the $\Delta D_{95\%}$ of the GTV, was considered the gold standard. This metric was deemed positive, and to represent a relevant dose deviation due to the simulated anatomical change, when it exceeded 4% (Brahme 1984), and negative otherwise. With this information, it could be established whether, for instance, a positive gamma fail rate result coincides with a positive $\Delta D_{95\%}$, constituting a true positive result. In a similar way, true negatives (both gamma fail rate and $\Delta D_{95\%}$ negative), false positives (gamma fail rate positive but $\Delta D_{95\%}$ negative) and false negatives (gamma fail rate negative but $\Delta D_{95\%}$ positive) could be determined.

Based on these definitions, confusion matrices containing the number of true positives, true negatives, false positives and false negatives can be created, from which the sensitivity (true positive rate) and specificity (true negative rate) can be calculated. ROC curves are a graphical display of the sensitivity and (1-specificity) over a range of thresholds of the diagnostic test value, which was the gamma fail rate in this work. Therefore, ROC curves were generated in MATLAB for each combination of portal dosimetry method and gamma criteria. For each ROC curve the area under the curve (AUC) was calculated, as this is a measure of test validity (Kumar and Indrayan 2011). Hence, the combination of portal dosimetry method and gamma criteria with the highest AUC was considered to be the one that could detect relevant anatomical changes most accurately.

To establish which gamma fail rate threshold maximizes the correct classification of relevant and irrelevant dose deviations due to anatomical changes, the optimal cut-off point on each ROC curve was determined. This point was taken as the point geometrically closest to (0,1), where both sensitivity and specificity are 1 (Kumar and Indrayan 2011). The threshold corresponding to this cut-off point was taken as the optimal gamma fail rate threshold.

3. Results

3.1. Anatomical change simulation

Figure 2 shows the DVH curves for each level of simulated tumor shift, tumor regression and pleural effusion for one patient.

In the case of simulated tumor shift, the DVH curves in figure 2(a) show that the larger the shift, the more the distinct corner at the top right is disappearing. This is due to the GTV shifting away from the high dose region, which means that the tumor volume receiving a high dose decreases. For this patient, there is first a slight improvement in the DVH curves before this effect occurs. This is because the dose is not homogeneously distributed over the GTV, due to nearby organs-at-risk (OARs) that needed to be avoided. By shifting the GTV, it is first moved into a slightly higher dose region, where it is more uniformly covered, before showing the

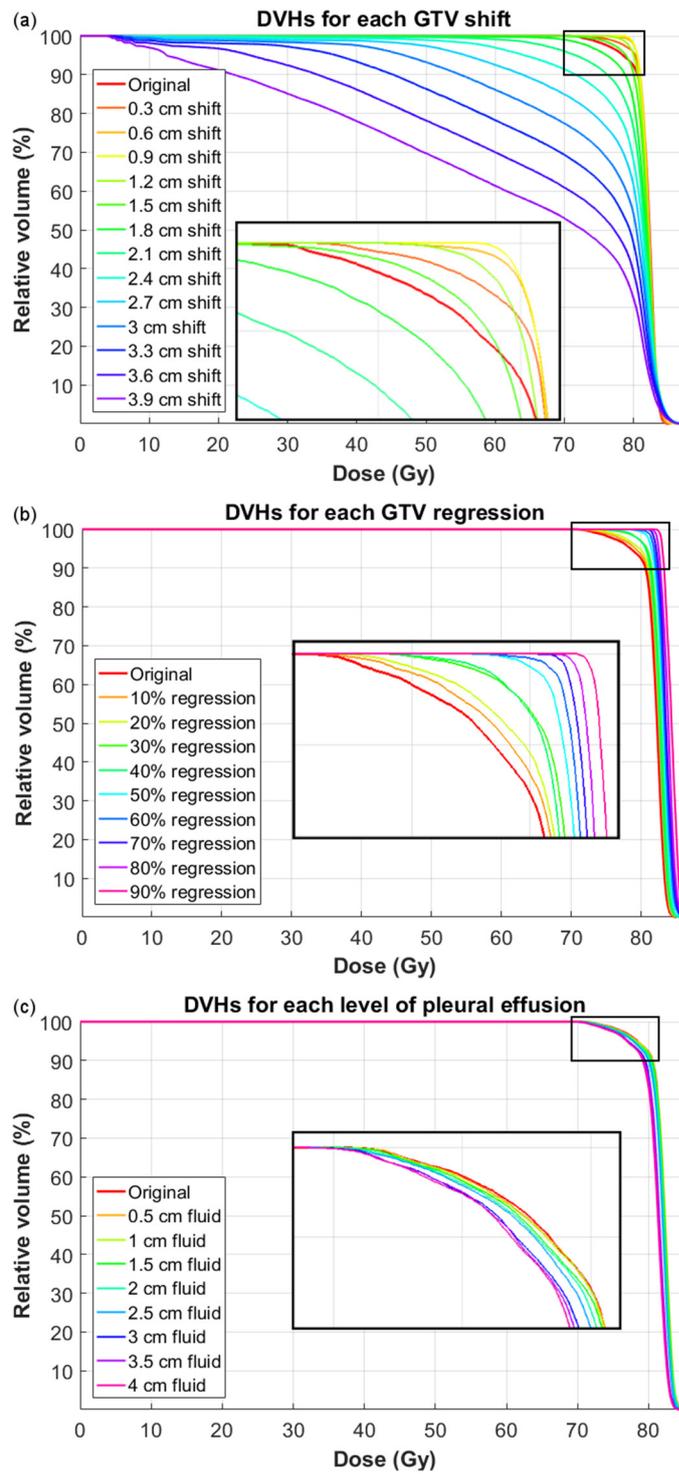


Figure 2. DVH curves for the GTV of one patient, for (a) tumor shift, (b) tumor regression and (c) pleural effusion.

expected decrease in volume receiving high dose. This occurs in three out of the six patients analyzed. For the other three patients the expected decrease in the volume receiving high dose is visible immediately. The changes in $D_{95\%}$ observed for tumor shift are generally large, and can even be up to 95% for the largest shifts and smaller tumor volumes.

The DVH curves corresponding to tumor regression in figure 2(b) show the opposite effect. There, the dose to the GTV seems to become higher with increasing tumor regression. This is because high density tumor tissue is replaced by low density lung tissue all around the shrinking tumor. Therefore, there is less attenuation from all sides of the tumor, which increases the dose to the GTV. For all patients, the same change in DVH curves is seen, albeit slightly larger for the three patients for which the GTV was not homogeneously covered in the original situation. For these three patients, as the tumor regression increases, a larger proportion of the relative tumor volume receives a high dose, as a larger part of the GTV is uniformly covered in the high dose region. The largest change in $D_{95\%}$ that occurs for tumor regression is approximately 6%.

In the DVHs for pleural effusion, shown in figure 2(c), it can be seen that the curves move slightly to lower doses with increasing level of pleural effusion. Like tumor shift, the volume of the GTV receiving high dose is decreasing. However, in this case this is due to the pleural effusion fluid attenuating the posterior beams more than lung tissue would, hence there is less dose reaching the tumor. Even though the beams do pass through the simulated pleural effusion volume in all patients (the half-arcs go from 180° to 0° , the full arcs from 180° back to 180°), the attenuation only happens in a small part of the 180° or 360° VMAT arc. In fact, for pleural effusion, no relevant changes (i.e. $>4\%$) in the $D_{95\%}$ are observed, with the largest change being 2%.

Figure 3 shows 3D bar charts for the number of patients for which each method can detect the simulated anatomical changes, for the various gamma criteria. It can be seen that in all cases, different methods can detect different levels of the simulated anatomical changes, and the gamma criteria that are used also influence this.

As is apparent from figure 3(a), the 2D-TI method can detect simulated tumor shift in most patients for tight gamma criteria, such as (1%, 1 mm), except for the smallest shifts. However, its performance decreases quickly for looser gamma criteria. 2D-TR shows this trend as well, although it is able to detect deviations in more patients than 2D-TI for the intermediate gamma criteria. 3D-TI performs best in detecting simulated tumor shift, but for looser gamma criteria it can only detect the larger shifts. From the total amount of simulated shifts for all patients and gamma criteria, which was 468, 2D-TI could detect 29.5%, 2D-TR detected 38.7% and 3D-TI outperformed the other methods by detecting 78.2%.

In figure 3(b), it can be seen that 3D-TI outperforms 2D-TI and 2D-TR in detecting tumor regression, for tight gamma criteria. For looser criteria none of the methods perform well, with 3D-TI not detecting the tumor regressions in any of the patients for the (5%, 3 mm) gamma criteria. For (1%, 1 mm), 2D-TR is able to detect deviations in more patients than 2D-TI, but for looser gamma criteria their performance is similar. Looking at the percentage of simulated regressions the different methods detected from a total of 324, 2D-TI, 2D-TR and 3D-TI detected 12.3%, 15.7% and 36.4%, respectively.

Figure 3(c) shows that for the tightest gamma criteria simulated pleural effusion can be detected in most patients, especially using 2D-TR or 3D-TI. For looser gamma criteria, 3D-TI cannot detect deviations in any of the patients, while 2D-TR performs better than 2D-TI in these cases. The total amount of simulated pleural effusion cases was 288, of which 2D-TI detected 18.8%, 2D-TR 31.6% and 3D-TI 19.4%.

Generally, except for pleural effusion, the simulated anatomical changes are large enough to cause significant deviations in the $D_{95\%}$. 3D-TI is able to detect errors in more cases than

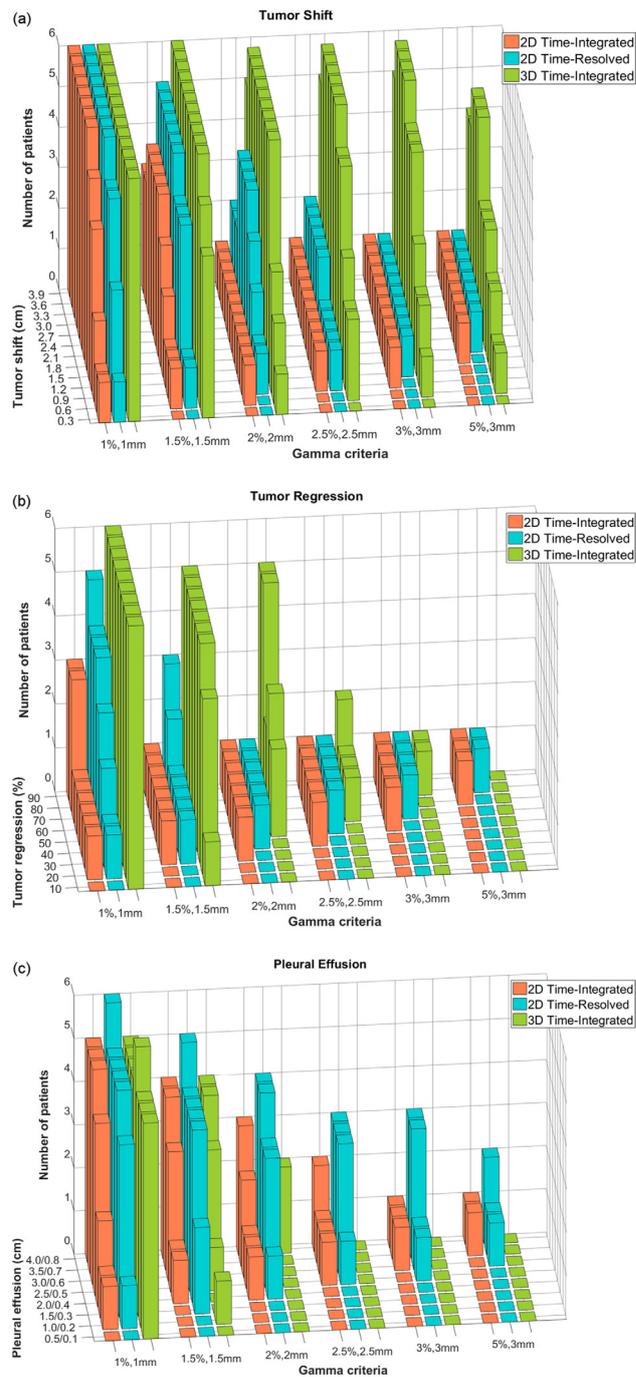


Figure 3. 3D bar charts of the number of patients for which the gamma fail rate exceeds the 10% threshold for each simulated anatomical change and for different gamma criteria, for (a) tumor shift, (b) tumor regression and (c) pleural effusion there were two options for the simulated levels: 0.5–4 cm or 0.1–1 cm. Both of these options are taken together in this graph, as the pleural effusion level relative to the tumor position is similar.

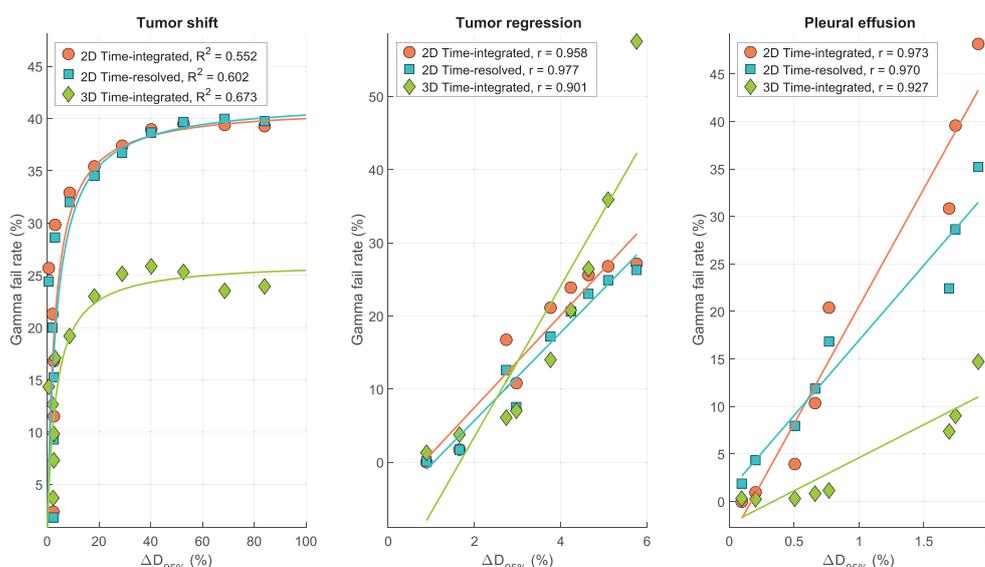


Figure 4. Correlation between $\Delta D_{95\%}$ and gamma fail rate for one individual patient, for gamma criteria of (2%, 2mm). For tumor shift the relationship was fitted to a saturation-growth curve.

2D-TI and 2D-TR. For tumor regression and pleural effusion, 3D-TI detects smaller anatomical changes when tighter gamma criteria are used, while for looser criteria the 2D methods are better.

3.2. Correlation between DVH metrics and gamma fail rate

As outlined previously, a quantitative relationship between the outcomes of the portal dosimetry methods (i.e. gamma fail rates) and DVH metrics should be established. Figure 4 shows the results of the correlation analysis within a single patient (the same patient for which the DVH curves are shown in figure 2) for one of the sets of gamma criteria evaluated. It can be seen that the correlation is generally strong, especially for tumor regression and pleural effusion. For tumor shift, there was a clear non-linear relationship in the gamma fail rate versus $\Delta D_{95\%}$ curve, as this curve starts to saturate when the GTV shifts outside the PTV. This effect is seen for all patients, with the onset of the saturation dependent on the magnitude of the shift needed to move the tumor out of the PTV, and the rate of saturation dependent on the tumor volume. Therefore, a non-linear fit was performed using a saturation-growth curve. For all three types of anatomical changes, similar trends were observed for the other gamma criteria.

An overview of the correlation coefficients for all patients on an individual level is provided in figure 5. Generally, there is a stronger relationship for tumor regression than for shift and pleural effusion. For all simulated anatomical changes, there are several cases in which the correlation coefficient or coefficient of determination could not be calculated. This occurs when the portal dosimetry method did not detect any errors, for instance with simulated pleural effusion, 3D-TI and gamma criteria of (5%, 3 mm). Generally, this occurred more frequently for looser gamma criteria. For most patients there are a number of cases in which there is a strong relation between the gamma fail rate and $\Delta D_{95\%}$. However, there are also cases for which the correlation is weak or even non-existent, such as for pleural effusion in patient 6.

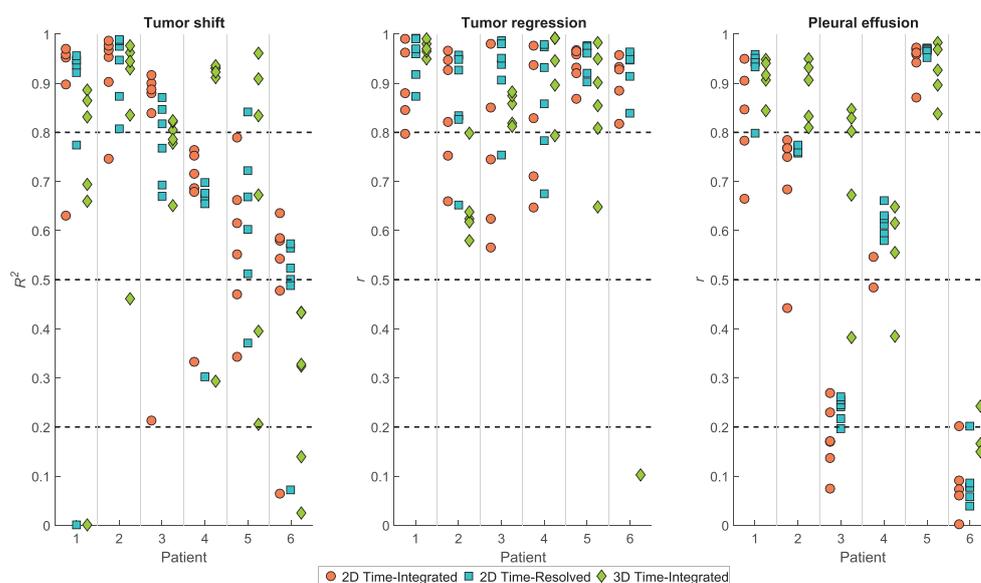


Figure 5. Correlation coefficients for all patients individually. Each marker represents the correlation coefficient r or coefficient of determination R^2 for a certain patient, portal dosimetry method and gamma criteria. If there are less than 6 markers (one for each set of gamma criteria) per method per patient, r or R^2 could not be calculated for all gamma criteria. The dashed lines indicate the regions in which the coefficients represent a non-existent, weak, moderate and strong linear relationship, as outlined in section 2.2. Patient 5 is the patient whose individual results are displayed in figure 4. Not displayed are the weak negative correlation coefficients for patient 6 in the case of tumor regression.

When looking at the six patients as a group, the relationship between the gamma fail rate and $\Delta D_{95\%}$ is less strong (figure 6). Especially in the case of tumor shift, the correlation is non-existent for 2D, and only moderate for 3D. For tumor regression, the correlation is also moderate. There seems to be a stronger correlation for pleural effusion. However, as mentioned previously, there are no significant deviations ($>4\%$) in the $D_{95\%}$ for this type of simulated anatomical change. When applying the other gamma criteria, these trends are equivalent.

3.3. Sensitivity and specificity analysis

The sensitivity and specificity analysis was performed for tumor shift and tumor regression only, since there were no significant deviations in the $D_{95\%}$ for pleural effusion. For this analysis, the data of all patients with respect to an anatomical change type is combined. A true positive result occurs when both the gamma fail rate and the $D_{95\%}$ deviate significantly, i.e. the gamma fail rate exceeds a certain threshold in the in-field region in 2D or in the GTV structure in 3D, and $\Delta D_{95\%}$ is larger than 4%.

Figure 7 shows the ROC curves and AUC values for simulated tumor shift. Based on the AUC values, when the gamma criteria are tight, 2D-TI performs slightly better than 2D-TR, but when the gamma criteria are looser, 2D-TR does much better. Moreover, the method that performs best is 3D-TI, particularly with gamma criteria of (2%, 2 mm) or (2.5%, 2.5 mm). Based on the optimal cut-off points on the ROC curves, the corresponding gamma fail rate thresholds for this best-performing method with these gamma criteria, are 19% and 12%,

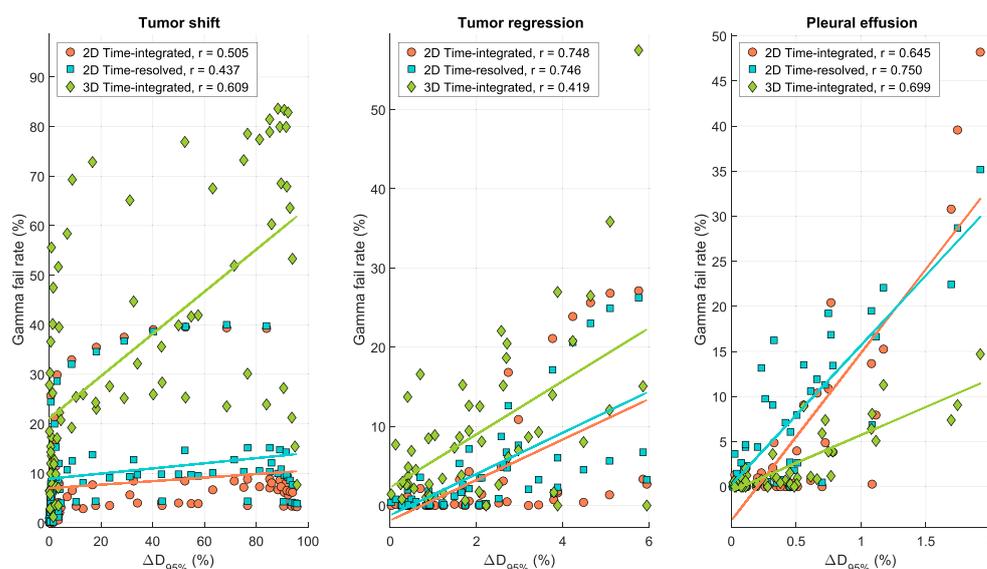


Figure 6. Correlation between $\Delta D_{95\%}$ and the gamma fail rate for the six patients as a group, for gamma criteria of (2%, 2 mm).

respectively. The confusion matrix for these optimal cases, displayed in table 2(a), shows the achievable frequencies of correct classifications of relevant simulated tumor shift. This confusion matrix applies both to 3D-TI with gamma criteria (2%, 2 mm) and fail rate threshold 19% and to 3D-TI with (2.5%, 2.5 mm) and 12%. The sensitivity and specificity are 93.6% and 71.0%, respectively.

A similar analysis for tumor regression, displayed in figure 8, reveals that 2D-TR has the most stable performance over all gamma criteria, i.e. its AUC changes the least for different gamma criteria. Especially for the (5%, 3 mm) criteria the performance of 2D-TI and 3D-TI is far worse than that of 2D-TR. This is also visible in the ROC curve for the (5%, 3 mm) criteria. With 2D-TI and 3D-TI for these criteria, there are very few cases for which the gamma fail rate is larger than 0. Thus, only a few thresholds can be set to obtain the ROC curves. The optimal combination of method, gamma criteria and gamma fail rate threshold, is 2D-TR with (1%, 1 mm) criteria, and a threshold of 18%. For this combination, for which the confusion matrix is shown in table 2(b), the sensitivity and specificity are 75.0% and 93.5%, respectively.

To compare with the commonly used (3%, 3 mm) gamma criteria and 10% gamma fail rate threshold, in table 3 the sensitivity and specificity for each method with these criteria and fail rate threshold are displayed. Except for 3D-TI and tumor shift, the sensitivity is low, meaning that there are many false negatives when these parameters are used. On the contrary, the specificity is high, so there are not many false positive cases. Especially for tumor regression there is a large difference in the sensitivity that can be achieved.

4. Discussion

It is evident that there is a difference in the ability of 2D-TI, 2D-TR and 3D-TI portal dosimetry to detect different simulated anatomical changes. Their performance also depends on the gamma criteria that are used, as more errors are detected with tighter gamma criteria.

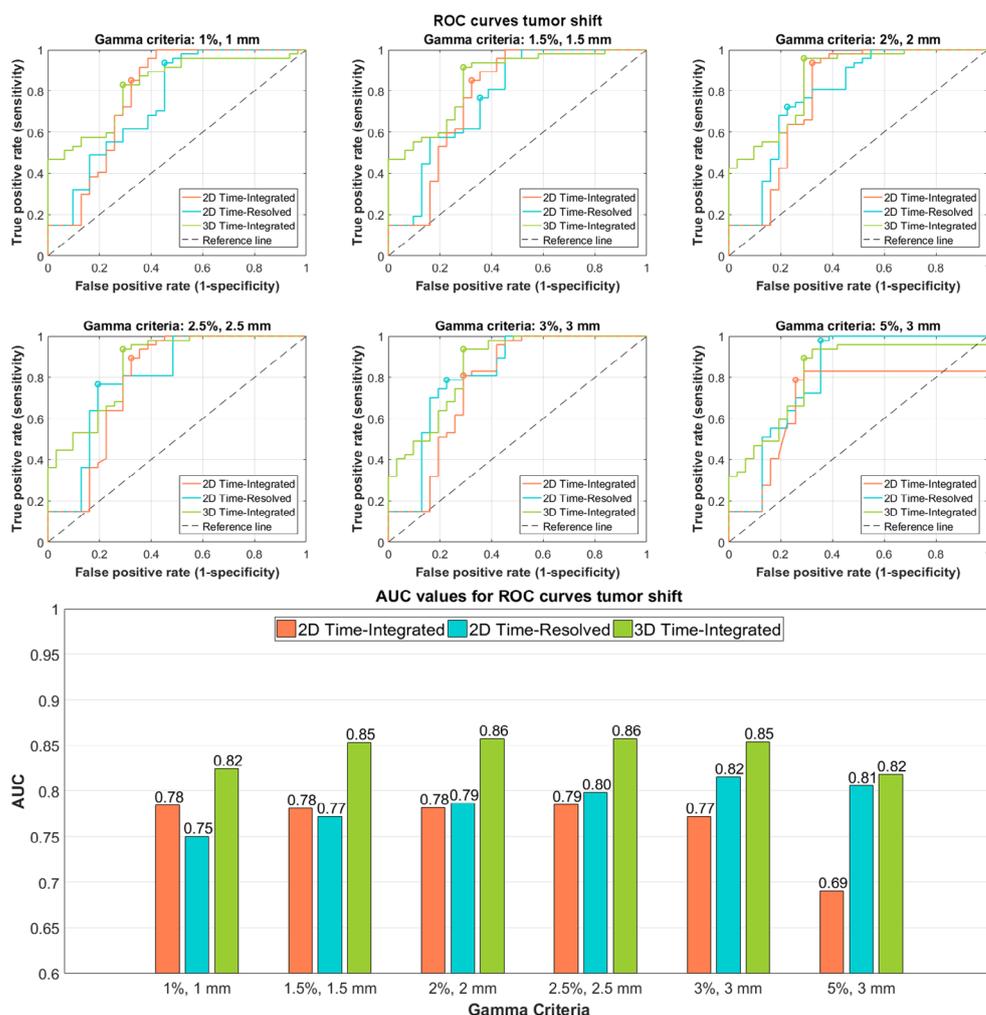


Figure 7. ROC curves and AUC values for tumor shift. The markers on the ROC curves represent the optimal cut-off point, i.e. the point on each curve closest to (0, 1).

Table 2. Confusion matrices (displaying the frequencies of positive and negative detection rates) for the optimal methods for detecting anatomical changes: (a) 3D-TI with gamma criteria (2%, 2 mm) and gamma fail rate threshold of 19%, or with gamma criteria (2.5%, 2.5 mm) and gamma fail rate threshold of 12% for tumor shift (their confusion matrices are exactly the same), and (b) 2D-TR with gamma criteria (1%, 1 mm) and gamma fail rate threshold of 18% for tumor regression.

(a)	Optimal confusion matrix	DVH analysis		(b)	Optimal confusion matrix	DVH analysis		
		Positive	Negative			Positive	Negative	
	Tumor shift				Tumor regression			
	Gamma analysis	Positive	44	9	Gamma analysis	Positive	6	3
		Negative	3	22		Negative	2	43

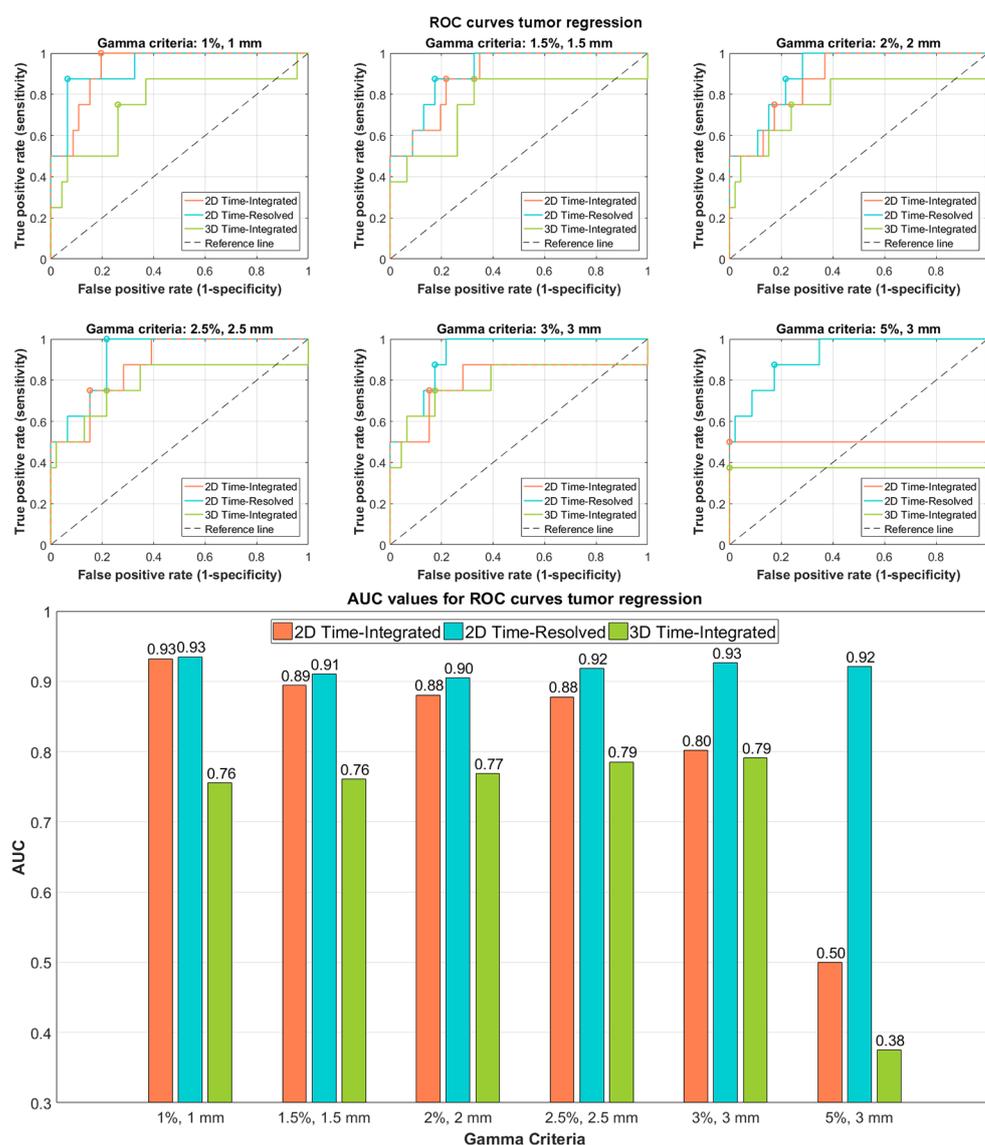


Figure 8. ROC curves and AUC values for tumor regression. The markers on the ROC curves represent the optimal cut-off point, i.e. the point closest to (0, 1).

Naturally, not all of the dose deviations detected with the portal dosimetry methods are clinically relevant, and it needs to be established which ones are, also taking the uncertainties of the methods into account when using them for actual measurements instead of simulations.

The most important result of this work is that there is a relationship at the patient level between the gamma fail rate and $\Delta D_{95\%}$, for many combinations of portal dosimetry methods and gamma criteria. However, at the cohort level the relationship is less strong, and in many cases non-existent. This implies that it will be more beneficial to relate portal dosimetry and DVH analysis on the patient level rather than trying to quantify this relationship for a group of patients, as has been done previously (Waghorn *et al* 2011, Persoon *et al* 2015a, Schyns *et al* 2016).

Table 3. The sensitivity and specificity for simulated anatomical changes (tumor shift and tumor regression) and each portal dosimetry method, for gamma criteria (3%, 3 mm) and a gamma fail rate threshold of 10%. The sensitivity and specificity of the optimal combinations of portal dosimetry method, gamma criteria and gamma fail rate threshold (see table 2) are displayed as well for comparison.

	Tumor shift				Tumor regression			
	2D-TI	2D-TR	3D-TI	Optimal	2D-TI	2D-TR	3D-TI	Optimal
Sensitivity (%)	14.9	14.9	91.5	93.6	50	50	25	75
Specificity (%)	87.1	90.3	71.0	71.0	91.8	100	100	93.5

Based on these results, in clinical practice, for each patient, gamma fail rate versus $\Delta D_{95\%}$ curves could be simulated for certain anatomical changes prior to treatment. These simulated curves could then serve to decide when the gamma fail rate measured during treatment exceeds a level that indicates a clinically relevant change in the $D_{95\%}$. As the relationship between the gamma fail rate and the $\Delta D_{95\%}$ is not the same for all anatomical changes, simulating a curve for each type of anatomical change might be necessary.

Although this work focused on the $D_{95\%}$, the correlation between other DVH metrics and the gamma fail rate on the patient level should be investigated as well. For instance, D_{mean} and $V_{95\%}$ change in a similar manner to the $D_{95\%}$ for the anatomical changes simulated in this study, and thus likely correlate with the gamma fail rate as well for individual patients. Other measures, such as dose conformity, could additionally be evaluated.

Furthermore, evaluating and relating DVH metrics for OARs to the gamma fail rate could provide additional insight, as the simulated anatomical changes in this study do not only influence the dose to the GTV, but also the dose to, for instance, the healthy lung tissue. The constraints for the OARs used during treatment planning are potentially violated due to anatomical changes. A thorough analysis of the effect of anatomical changes on the OARs could aid in developing criteria to decide when to adapt a treatment plan.

The sensitivity and specificity analyses provide information on which combination of portal dosimetry method, gamma criteria and gamma fail rate threshold to use to detect certain types of anatomical changes with optimal sensitivity and specificity. It should be noted that the value of $\Delta D_{95\%}$ that is considered significant influences the outcomes of these analyses. Moreover, in this work the optimal threshold was found in a geometrical way from the ROC curve, while in clinical practice a pre-specified sensitivity or specificity might be preferable. However, the results reveal that instead of using only one set of gamma criteria and a single fail rate threshold, different methods with different gamma criteria and thresholds are needed to optimally detect different types of relevant errors. As was clear from previous work as well (Yan *et al* 2009, Stasi *et al* 2012, Heilemann *et al* 2013), the most commonly used (3%, 3 mm) criteria are not sensitive enough to detect many significant anatomical changes (table 3). Using these criteria and a 10% gamma fail rate threshold does not optimize the trade-off between sensitivity and specificity.

In clinical practice, more uncertainties will be introduced than in this simulation study. Breathing motion, patient positioning, timing and setup are all factors that influence the ability of the portal dosimetry methods to detect relevant anatomical changes. Although, for instance, patient setup errors will be minimized before dose delivery by means of image guided radiotherapy, the gamma fail rate thresholds used in clinical practice likely need to be higher to account for these uncertainties. For the TR gamma analysis, the optimal time-to-agreement and time of interest (Podesta *et al* 2014a, 2014b) need to be determined as well in a clinical setting, as in reality the synchronization in time will not be perfect.

An additional difficulty is that more than one anatomical change can occur simultaneously, e.g. pleural effusion may cause a tumor shift (Møller *et al* 2014), and these may not be easily separable based on gamma analysis results. Furthermore, other anatomical changes can appear in lung cancer patients. Tumor progression is one example, although occurring infrequently (van Zwienen *et al* 2008) and therefore not considered in this work. Atelectasis is seen more often, but cannot be simulated consistently due to its unpredictable nature. As atelectasis is likely to cause tumor shift (Persoon *et al* 2013), it might be possible to detect this indirect anatomical change instead.

Other methods for evaluating anatomical changes exist, such as visual inspection of CBCT images or tumor tracking (Shimizu *et al* 2001, Belderbos *et al* 2007). However, these do not allow direct evaluation of the dose deviations caused by anatomical changes. Dose recalculation based on CBCT images does allow assessment of dose changes due to changing anatomy, but requires daily CBCT imaging. As was shown, large amounts of pleural effusion, which might be a cause for plan adaptation solely based on visual inspection, do not necessarily cause large deviations in the dose distribution. Hence, it is important to combine in-room imaging techniques with portal dosimetry, for instance by recalculating the 3D dose (based on PDI) on CBCT images. This can also aid in determining which anatomical change occurred.

5. Conclusion

This simulation study investigated the ability of 2D-TI, 2D-TR and 3D-TI portal dosimetry to detect dosimetric consequences of anatomical changes in lung cancer patients. This work confirms that transit portal dosimetry can be used to detect (simulated) anatomical changes in lung cancer patients, provided they are of a certain magnitude. The correlation between the gamma fail rate and DVH metrics at the patient level can be useful for creating decision protocols for adaptive radiotherapy. It may be possible by individual simulation to establish a relationship between DVH metrics and gamma fail rates before treatment. This was not possible before, as on the group level there is no correlation. Error-type specific optimal combinations of a portal dosimetry method, gamma criteria and gamma fail rate threshold should be used to determine if an anatomical change occurred and whether it is clinically relevant. In future work, this research should be extended to include, for instance, different tumor sites, DVH metrics, dose to OARs and 3D-TR portal dosimetry, which may be able to detect changes that are hidden by the 3D-TI method due to the integration. Furthermore, the uncertainties of the portal dosimetry methods need to be quantified before clinical decision models can be based on their outcomes. To conclude, this work shows that clinical decision models for lung cancer patients should be developed on an individual basis, and that a one-size-fits-all approach with respect to portal dosimetry method, gamma criteria and gamma fail rate is not suitable for detecting different types of anatomical changes.

Acknowledgments

Varian Medical Systems provided the funding for this study (project: TimePort-III).

References

- Bedford J L, Hanson I M and Hansen V N 2014 Portal dosimetry for VMAT using integrated images obtained during treatment *Med. Phys.* **41** 021725

- Belderbos J, van Beek S, van Kranen S, Rasch C, van Herk M and Sonke J 2007 Anatomical changes during radiotherapy of lung cancer patients *Int. J. Radiat. Oncol. Biol. Phys.* **69** S508–S9
- Bortfeld T 2006 IMRT: a review and preview *Phys. Med. Biol.* **51** R363–79
- Brahme A 1984 Dosimetric precision requirements in radiation therapy *Acta Radiol. Oncol.* **23** 379–91
- Britton K R, Starkschall G, Tucker S L, Pan T, Nelson C, Chang J Y, Cox J D, Mohan R and Komaki R 2007 Assessment of gross tumor volume regression and motion changes during radiotherapy for non-small-cell lung cancer as measured by four-dimensional computed tomography *Int. J. Radiat. Oncol. Biol. Phys.* **68** 1036–46
- Fraass B, Doppke K, Hunt M, Kutcher G, Starkschall G, Stern R and Van Dyke J 1998 American association of physicists in medicine radiation therapy committee task group 53: quality assurance for clinical radiotherapy treatment planning *Med. Phys.* **25** 1773–829
- Heilemann G, Poppe B and Laub W 2013 On the sensitivity of common gamma-index evaluation methods to MLC misalignments in Rapidarc quality assurance *Med. Phys.* **40** 031702
- Kumar R and Indrayan A 2011 Receiver operating characteristic (ROC) curve for medical researchers *Indian Pediatr.* **48** 277–87
- Kwint M, Conijn S, Schaake E, Kneijens J, Rossi M, Remeijer P, Sonke J-J and Belderbos J 2014 Intra thoracic anatomical changes in lung cancer patients during the course of radiotherapy *Radiother. Oncol.* **113** 392–7
- Liu B, Adamson J, Rodrigues A, Zhou F, Yin F-F and Wu Q 2013 A novel technique for VMAT QA with EPID in cine mode on a Varian TrueBeam linac *Phys. Med. Biol.* **58** 6683
- Low D A and Dempsey J F 2003 Evaluation of the gamma dose distribution comparison method *Med. Phys.* **30** 2455–64
- Low D A, Harms W B, Mutic S and Purdy J A 1998 A technique for the quantitative evaluation of dose distributions *Med. Phys.* **25** 656–61
- Low D A, Morele D, Chow P, Dou T H and Ju T 2013 Does the γ dose distribution comparison technique default to the distance to agreement test in clinical dose distributions? *Med. Phys.* **40** 071722
- Mans A, Wendling M, McDermott L N, Sonke J-J, Tielenburg R, Vijlbrief R, Mijnheer B, van Herk M and Stroom J C 2010 Catching errors with *in vivo* EPID dosimetry *Med. Phys.* **37** 2638–44
- McDermott L N, Wendling M, Sonke J-J, van Herk M and Mijnheer B J 2006 Anatomy changes in radiotherapy detected using portal imaging *Radiother. Oncol.* **79** 211–7
- Møller D S, Khalil A A, Knap M M and Hoffmann L 2014 Adaptive radiotherapy of lung cancer patients with pleural effusion or atelectasis *Radiother. Oncol.* **110** 517–22
- Nijsten S M J J G, Mijnheer B J, Dekker A L A J, Lambin P and Minken A W H 2007 Routine individualised patient dosimetry using electronic portal imaging devices *Radiother. Oncol.* **83** 65–75
- Persoon L C, Nijsten S M, Wilbrink F J, Podesta M, Snaith J A, Lustberg T, van Elmpt W J, van Gils F and Verhaegen F 2012 Interfractional trend analysis of dose differences based on 2D transit portal dosimetry *Phys. Med. Biol.* **57** 6445–58
- Persoon L C, Podesta M, Hoffmann L, Sanizadeh A, Schyns L E, de Ruiter B M, Nijsten S M, Muren L P, Troost E G and Verhaegen F 2015a Is integrated transit planar portal dosimetry able to detect geometric changes in lung cancer patients treated with volumetric modulated arc therapy? *Acta Oncol.* **54** 1501–7
- Persoon L C, Podesta M, Nijsten S M, Troost E G and Verhaegen F 2015b Time-resolved versus integrated transit planar dosimetry for volumetric modulated arc therapy: patient-specific dose differences during treatment, a proof of principle *Technol. Cancer Res. Treat.* **15** NP79–87
- Persoon L C G G, Egelmeier A G T M, Öllers M C, Nijsten S M J J G, Troost E G C and Verhaegen F 2013 First clinical results of adaptive radiotherapy based on 3D portal dosimetry for lung cancer patients with atelectasis treated with volumetric-modulated arc therapy (VMAT) *Acta Oncol.* **52** 1484–9
- Podesta M, Nijsten S M, Persoon L C, Scheib S G, Baltas C and Verhaegen F 2014a Time dependent pre-treatment EPID dosimetry for standard and FFF VMAT *Phys. Med. Biol.* **59** 4749–68
- Podesta M, Persoon L C and Verhaegen F 2014b A novel time dependent gamma evaluation function for dynamic 2D and 3D dose distributions *Phys. Med. Biol.* **59** 5973–85
- Schyns L E, Persoon L C, Podesta M, van Elmpt W J and Verhaegen F 2016 Time-resolved versus time-integrated portal dosimetry: the role of an object's position with respect to the isocenter in volumetric modulated arc therapy *Phys. Med. Biol.* **61** 3969–84
- Shimizu S, Shirato H, Ogura S, Akita-Dosaka H, Kitamura K, Nishioka T, Kagei K, Nishimura M and Miyasaka K 2001 Detection of lung tumor movement in real-time tumor-tracking radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **51** 304–10
- Spreeuw H, Rozendaal R, Olaciregui-Ruiz I, Gonzalez P, Mans A, Mijnheer B and van Herk M 2016 Online 3D EPID-based dose verification: proof of concept *Med. Phys.* **43** 3969

- Stasi M, Bresciani S, Miranti A, Maggio A, Sapino V and Gabriele P 2012 Pretreatment patient-specific IMRT quality assurance: a correlation study between gamma index and patient clinical dose volume histogram *Med. Phys.* **39** 7626–34
- Teoh M, Clark C, Wood K, Whitaker S and Nisbet A 2014 Volumetric modulated arc therapy: a review of current literature and clinical use in practice *Br. J. Radiol.* **84** 967–96
- van Elmpt W, McDermott L, Nijsten S, Wendling M, Lambin P and Mijnheer B 2008 A literature review of electronic portal imaging for radiotherapy dosimetry *Radiother. Oncol.* **88** 289–309
- van Elmpt W J, Nijsten S M, Mijnheer B J and Minken A W 2005 Experimental verification of a portal dose prediction model *Med. Phys.* **32** 2805–18
- van Elmpt W J, Nijsten S M, Schiffeleers R F, Dekker A L, Mijnheer B J, Lambin P and Minken A W 2006 A Monte Carlo based three-dimensional dose reconstruction method derived from portal dose images *Med. Phys.* **33** 2426–34
- van Zwienen M, van Beek S, Belderbos J, van Kranen S, Rasch C, van Herk M and Sonke J 2008 Anatomical changes during radiotherapy of lung cancer patients *Int. J. Radiat. Oncol. Biol. Phys.* **72** S111
- Waghorn B J, Meeks S L and Langen K M 2011 Analyzing the impact of intrafraction motion: correlation of different dose metrics with changes in target $D_{95\%}$ *Med. Phys.* **38** 4505–11
- Woodruff H C, Fuangrod T, Rowshanfarzad P, McCurdy B M C and Greer P B 2013 Gantry-angle resolved VMAT pretreatment verification using EPID image prediction *Med. Phys.* **40** 081715
- Yan D, Vicini F, Wong J and Martinez A 1997 Adaptive radiation therapy *Phys. Med. Biol.* **42** 123
- Yan G, Liu C, Simon T A, Peng L C, Fox C and Li J G 2009 On the sensitivity of patient-specific IMRT QA to MLC positioning errors *J. Appl. Clin. Med. Phys.* **10** 2915
- Zou K H, Tuncali K and Silverman S G 2003 Correlation and simple linear regression *Radiology* **227** 617–22